Int. J. Advance Soft Compu. Appl, Vol. 9, No. 3, Nov 2017
ISSN 2074-8523

A Bibliometric Mapping of Malaysian Publication using Co-Word Analysis

Nurul Mardhiah Azura Md Nadzar [1], Aryati Bakri [2], and Roliana Ibrahim[3]

[1,2,3] Faculty of Computing, Universiti Teknologi Malaysia
e-mail: [1]mardhiahazura@gmail.com, [2]aryati@utm.my, [3]roliana@utm.my

Abstract

Bibliometric maps or science maps are figurative representation of scientific field or organization in which elements of the map are associated with topic or themes. In this study, the method that has been chosen to perform bibliometric mapping is co-words analysis. Main objective for this study is to identify the research progress that been achieved by previous researchers and future direction which has the potential to be discovered by younger researchers. Set of data been downloaded from Scopus database. In this study total of 2360 successfully downloaded which has been sorted the highest cited articles and all subject area is selected from year 1995 to 2015. By using Vosviewer, title and abstract from corpus is manipulated in order to produce cluster of co-words analysis. As for result, there are 5 cluster words being produce in the software. Three of clusters are related to medicine, pharmacy and clinical namely as cluster 1, cluster 4 and cluster 5. In cluster 2 is related to agricultural, environment and forestry whereas the last cluster, cluster 3 is more relate to children development, adult health and health in developing countries. Highest and low frequency word also reported in this study. Highest frequency can be considered as the most research area that become researcher's choice in this past 20 years. Whereas for low frequency words is the potential words that can be a guidance for young researchers to choose which research area or domain that they can gain the interest and do the next exploration.

Keywords: *bibliometric mapping, co-words analysis, publication, cluster, research domain*

1 Introduction

Bibliometric mapping have been around for quite a while. Nonetheless, the number of paper remain low when it comes to identify the progress for research publication as total regardless the subject area or domain. There are number of publication which reviewed and studied a bibliometric mapping in certain domain such as computer intelligence (Eck and Waltman, 2007a) , education (Assefa and Rorissa, 2013), drinking water research (Fu *et al.*, 2013), medicine (Bazm *et al.*, 2016) and others (Lee and Jeong, 2008) but, just a few study been made to identify the progress of researcher's publication which happens in the certain country (Gazni *et al.*, 2012). One of main reason on bibliometric mapping is to identify the nature relationship and structure of the communities and it also knows as 'science map' (Buter *et al.*, 2006; Eck *et al.*, 2010) or landscape of science (Noyons, 1999a).

Number of publication are increasing and collaboration between countries as well as with research centers are keep on carried out by research universities. Undoubtedly, new discoveries are always being center of attention as one of greatest achievement by researchers and scientist. Although it is undeniable by the establishment of research universities and research centers, publication and research are vigorously done by these parties. Yet, we didn't know which area and research being center of attention and actively done by Malaysian's researchers. Therefore, in order to identify the progress and main focus of publication output, one of approach to get the overview is by perform the visualization bibliometric element namely as bibliometric mapping. Thus, it become the main objective for this study is to identify the research progress that been achieved by previous researchers and future direction which has the potential to be discovered by younger researchers. Since there are no study been made before in bibliometric mapping for Malaysian's researchers, thus, this is the first study to do the exploration what have been made by previous researchers as well as the domain research that become a choice. The exploration is by manipulated bibliometric element namely as title and abstract which consist in publication. It is important to discover this bibliometric map in order to make sure the research and development are keep on evolving and continue. It also can give an idea to researchers to seek the opportunities in explore the new topics or research area which has the potential to become one of their domain of study.

2 Related work

Bibliometric maps or science maps are figurative representation of scientific field or organization in which elements of the map are associated with topic or themes (Assefa and Rorissa, 2013). It also can be considered as a bridge to related research topic (Buter *et al.*, 2006; Glenisson *et al.*, 2005) or a graphical summaries of sets of papers, which is either it is based on sets of citation data

which normally called as co-citation maps (Egghe and Rousseau, 2002; Small, 1973), words of phrases which normally called as co-words analysis (Leydesdorff, 1996; Leydesdorff and Welbers, 2011) or some other bibliometric element (Buter *et al.*, 2006).

Bibliometric mapping is visualized multiple item which called nodes and has the underlying relationships between one nodes to another which called edges (Bazm *et al.*, 2016). The relationship edges comes from either authors, journals, abstract, titles or keywords which occurring in research papers (Assefa and Rorissa, 2013). Among the advantages performs bibliometric mapping is to get an idea and valuable insight on what is actually happening in certain area such as which domain is currently generate more papers or articles, which current issues that get the attention of current and previous researchers or how well the publication been made by the researchers (Muñoz-Leiva *et al.*, 2011). This valuable insight is can give such a good overview for management team to know the progress of their researchers and for younger researcher to get the surface information and for them to choose which path to choose in order to stamp their name in certain domain area (Noyons, 1999b; van Eck *et al.*, 2010).

There are a few type of bibliometric mapping namely as co-citation or bibliographic coupling (Boyack and Klavans, 2010), co-words (Igami *et al.*, 2014), co-authorship (Huang *et al.*, 2015), and others. Co-citation or bibliographic coupling is the condition where two documents is an item of reference used by two document. The strength of for this type of analysis is the number of reference which have in two documents in commons (Boyack and Klavans, 2010). Co-words analysis is the condition of the occurrences of keywords which extracted from the corpus's title, abstract or even full articles (Li and Chu, 2016) whereas as for co-authorship normally applied when two or more author collaborate with each other to produce some papers (Yan and Ding, 2012). Among all type of bibliometric mapping, selected type which is implemented in this paper is co-words analysis in order to produce the visualization of research topic and potential research area.

Output generated from bibliometric mapping normally in a form of cluster of a group (Andersen *et al.*, 2015). Group of important item which has higher frequency from the others which doesn't gain any high value (Bazm *et al.*, 2016). The most high frequency will lead the cluster and another related which has the strong connection with the main item is located in the same group (Bazm *et al.*, 2016). From this cluster, we can see the relationship between clusters by observe the inter-connection and intra-connection between items within the cluster (Andersen *et al.*, 2015). From generated cluster, more valuable information can be interpreted which can help number of parties such as researchers which have the interest in certain area or management team which has the intention to get the knowledge about the publication or author or another element in bibliometric or even for any funding bodies to have the overview on what's in going on in certain area research environment (Andersen *et al.*, 2015).

In certain way, mapping and clustering is complement each other (Assefa and Rorissa, 2013). Cluster analysis involves number of different algorithms which has the main purpose to detect natural division of networks into groups namely as cluster that based on similarity and minimize inter-cluster similarity (Bazm *et al.*, 2016). Once the relationship or distance between words were identified or establish, the cluster of words will be generated in order to represent the group of significance words which has the relationship between one and another. The close distance between words means the strong relationship and importance connection between each other. In one certain map, there are few clusters are generated and there are certainly some connection between clusters. Therefore, in bibliometric map, it is useful tools to study the structure and the dynamics of scientific research.

2.2 Co-words analysis

In co – words analysis, it is often accompanied by mapping and visualization of significant words, terms or phrases which to understand the structure of a bibliometric network, including the identification of major topics in a domain, the relationship between the topic and the diffusion of ideas over time (Dehdarirad *et al.*, 2014; Liu *et al.*, 2014; Viedma-Del-Jesus *et al.*, 2011). In bibliometric mapping, the frequent tools to analyze co-words analysis is either title or subtitle, author's keywords, journal's title, abstract or full articles (Eck and Waltman, 2007b; Heersmink *et al.*, 2011; Waltman and Eck, 2012). There are several technique which been used by previous researchers in order to retrieve certain frequency words or to calculate the weight of words which embedded in the abstract or full articles. Among all techniques frequently used by researchers are is multidimensional scaling (MDS) and visualization of similarities (Vos). Both MDS and VOS are basically use same approach which is multidimensional scaling (Eck *et al.*, 2010). Aim of MDS is to locate items in a low dimensional space in such a way that the distance between any two items reflects the similarity or relatedness of the items as accurately as possible (Eck *et al.*, 2010). The stronger the relation between two items, the smaller the distance between the items. Comparison between MDS and Vos is Vos is much more robust compared to MDS (Eck and Waltman, 2009; Eck *et al.*, 2010). Vos capable to produce a large complex map and able to process large amount of data as well as capable to produce clusters where each of cluster contain high similarities between nodes (Eck and Waltman, 2007c, 2011; Eck and Waltman., 2010)

One of benefit in co-words is a very useful tools which allows researchers as well as the management team to get the overview and identify the key patterns and trends which has happened and to get the idea the future exploration within the area. As overview, co words analysis is a content analysis that effective in mapping the strength of association between information items in textual data. Co-words known as maps (Carley, 1997), term mapping (Rokaya *et al.*, 2008),

bibliometric mapping(van Eck *et al.*, 2010), keywords map(Chen *et al.*, 2015), text mapping(Eck and Waltman, 2011), networks of words (Danowski, 1993), networks of concepts (Popping, 2000) or semantic network (Lehmann, 1992). Whichever term been used, the purpose in performing the co-words analysis is pretty much the same which is to analyze the text.

Nevertheless, to retrieve full articles seems the last choice due to limitation of library's subscription database, therefore, most of research used combination of title with the full abstract or title, subtitle and full abstract. The advantage exploit the abstract is normally abstract is a brief summary of a research articles, thesis, review, conference proceeding or any in-depth analysis of a particular subject and frequently used to help the interest reader quickly learn the purpose of papers (Ding *et al.*, 2001; Glenisson *et al.*, 2005; Janssens *et al.*, 2006). Therefore, there might be a high percentage on words repeating in abstract and full articles although some journal or publisher put a limitation words in abstract.

There are so specific definition about co –words analysis but most of the definition from other researcher is the content connection or communication in language which can be modeled as a networks of words and definitely has the relationship between each other (Callon *et al.*, 1991; Ding *et al.*, 2001; Janssens *et al.*, 2006) . This content analysis was claimed to be one of powerful technique for discovering and describing the interaction between different fields in scientific research (Bailón-Moreno *et al.*, 2006; Callon *et al.*, 1991; Leydesdorff and Zhou, 2008; Muñoz-Leiva *et al.*, 2011). Although there are study been made by other researchers by manipulate other language such as Chinese character, there is indeed a need the thesaurus for system or application to read and interpreted each characters (Gu and Li, 2015; Leydesdorff and Zhou, 2008).

There are huge benefits which was proven by previous researchers and has a huge impact in publication and research environment. One of the advantages is its help to plan or act as an aid in determine the research agenda as well as in evaluating the state of scientific field (Ding *et al.*, 2001; Janssens *et al.*, 2006). Co-words were tested and were validated that its map worked as reasonable representation of cognitive and social relation of the research field that have been studied by researchers (Healey, 1986; Peters and van Raan, 1993). Co words can act as the second opinion besides relies on expert's opinion only (Heersmink *et al.*, 2011).

This is not disclaim any expert's or senior's opinion or judgement in determine the direction of certain research as we still needs human experts thoughts and ideas. But then, sometimes human judgement is limited due to our limitations capabilities in remembering and discovering our surrounding and also our research interest. Therefore, co-words become one of medium to help between experts, seniors with other researchers to do more exploration in certain research area (Coulter *et al.*, 1998; Ding *et al.*, 2001; Looze and Lemarié, 2006). Most conflict happens when it comes to sub-subject area, or new research area in certain subject area that quite new and isn't mature enough to be discovered.

Certainly, researchers needs an aid and support to helps them to move or climb from novice stage to intermediate stage and so on. By observes and perceives words that being produces in co-words analysis, it might help them to get a little overview to do next steps.

Another advantages using the co-words analysis is has the capabilities to identify the research topic with referring to the statistical analysis that produces together with the occurrences words with only manipulate the previous publication that being published throughout the years (van Eck *et al.*, 2010; Viedma-Del-Jesus *et al.*, 2011). Previous publication which gain a lot of citation by other researchers certainly give an impact to the research environment as well as their readers, therefore, these document might has hidden information that need to be discovered by other researchers (Cho and Khang, 2006). The evolution from previous to current publication definitely shows huge discovery in the scientific field (Viedma-Del-Jesus *et al.*, 2011). Some of words might already saturated but some of them might has space to do the investigation or some of them can be manipulate to become another interesting discovery. Thus, to gain the insight words that generated from publication is important and can give huge opportunities to researches and field to growth and expend.

Some of co-words research are performed without any combination from other bibliometric analysis, but some of research did the combination such as with citation analysis or co-citation. The combination of indicator is to make sure the relatedness between publications (Waltman and Eck, 2012). Most common combination was between citation analysis with abstract and title (Heersmink *et al.*, 2011). Citation analysis normally shows the frequent number of publication being cited by other (Heersmink *et al.*, 2011; Janssens *et al.*, 2006). The higher numbers of citation, shows that publication did some impact to the research environment, therefore, it might have the hidden information that have the potential to be discovered by others. Another reason on selected the higher number of citation is typically the publication comes from experts or well-known researchers (Glänzel and Schubert, 2003). Usually research from the experts gain a lot of attention by others especially younger researcher in order to develop their name in research world. Certainly, experts has a lot more experience in performing the research and publication and the idea from the publication are definitely important and precious, therefore, it is significance to select the publication which has higher number of citation(Braam *et al.*, 1991; Glänzel and Schubert, 2003). Thus, in this study, before any retrieving activities being perform, all document are sorted according from higher number of citation to the small numbers of citation. This is to make sure that all document will be analyze precisely and all occurrences words being developed from the selected system are much more beneficial to others

With huge benefits can gain from co-words analysis, it is significant to implement in all related field in order to see and identify the progress of their research and it can give huge advantage in a field to do more investigation for new discovery.

There are numbers of system or application that been invented by researchers in determined words in the publication such as VOSVIEWER (Tahamtan *et al.*) (Eck and Waltman, 2009), LEXMAPPE (Ding *et al.*, 2001), Content Analysis and Information Retrieval (CAIR)(Coulter *et al.*, 1998), NPtools (Ding *et al.*, 2001) or other related software that capable to extract occurrences words from text. These systems or application has the skills to retrieved words from the text and has the capabilities to identify the most occurrences words in the publication. Most of system were developed as standalone system which means all systems need to be downloaded, installed and analyzed on personal workstation. This is because of most system are able process huge number of publication, only a few system from earlier era has the limitation in processing numbers of publication. As there are a lot of application exists in performing co-words analysis, one of the application namely as VosViewer is elected in this study. VosViewer or visualization of similarities was development by previous researchers namely as Nees Jan van Eck and Ludo Waltman (Eck and Waltman, 2011; Eck and Waltman., 2010). It is a computer program that was developed to creating, visualizing and exploring the bibliometric map. The program isn't available for collection data for web of science only, but it also compatible with data from Scopus and PubMed which specific database for medicine. This application can br used to analyzing all kind of bibliometric network data, for instance citation relations between publications or journals, collaboration relations between researchers, and co-occurrence relations between scientific terms and has the capabilities to handle thousands of data (Eck and Waltman, 2011).

2.2.1 Vosviewer

Basically VosViewer is the application which apply the concept of text mining. Text mining or knowledge discoveries (Feldman and Dagan, 1995) is a process of digging up the information from the textual database, extracting the interesting and non-trivial pattern or knowledge (Tan, 1990). This concept already exists quite sometimes and it does gives some impact to the research world with any hidden information from the text database can be brought out from manipulate the collection of text such as from text book (Berry, 2004), collection of publication(Gu and Li, 2015) , social media (He *et al.*, 2013), phone messages (Vlieger and Leydesdorff, 2011) and others that were textual based. The idea of created and developed this application is to provide the comprehensive graphical presentation for bibliometric data which previously, most of the application has limitation in viewing the huge number of text data (Eck and Waltman., 2010). Most of previous application only capable to own simple visualization with limited visual or graphical presentation (Ahlgren *et al.*, 2003; Richard Klavans and Boyack, 2006). Therefore, VosViewer has the intention to provide the graphical representation of large bibliometric maps can be much enhanced by means of, for example, zoom functionality, special labeling algorithms, and

density metaphors (Eck and Waltman., 2010). This kind of functionality is not incorporated into the computer programs that are commonly used by bibliometric researchers. This program pays special attention to the graphical representation of bibliometric maps.

In VosViewer, there are a few text mining functionality that provides term map based on corpus of document. The main component of Vosviewer are the mapping technique used for constructing map and the viewer used for examining maps. VOSviewer is the result of integrating the VOS mapping technique and the improved viewer software. VOSviewer can for example be used to view maps that have been constructed using a multidimensional scaling program such as the PROXSCAL program in SPSS. Another more technical thing to note is that VOSviewer has been written entirely in the Java programming language. Because of this, VOSviewer can be used on almost any hardware and operating system platform. It can also be started directly from a web page on the internet.

There are two types of bibliometric map that Vosviewer is intended to be used which are distance-based maps and graph-based maps. Distance-based maps are maps in which the distance between two items indicates the strength of the relation between the items. A smaller distance generally indicates a stronger relation. In most cases, items tend to be distributed quite unevenly in distance-based maps. On the one hand this makes it easy to identify clusters of related items,but on the other hand this sometimes makes it difficult to label all the items in a map without having labels that overlap each other. Whereas for graph-based map are maps in which the distance between two items need not indicate the strength of the relation between the items. Instead, lines are drawn between items to indicate relations. Items are usually distributed in a fairly uniform way in graph-based maps. This has the advantage that it is relatively easy to avoid the problem of overlapping labels. The disadvantage of graph-based maps compared with distance-based maps is that it typically is more difficult to see the strength of the relation between two items. Clusters of related items are also more difficult to detect.

VOSviewer can be used to view any distance-based map, regardless of the mapping technique with which the map has been constructed. One can use VOSviewer to view multidimensional scaling maps produced using statistical packages such as SAS, SPSS, and R, but one can also use VOSviewer to view maps constructed using less common techniques such as VxOrd. Because the VOS mapping technique shows a very good performance (Van Eck et al., 2008a), this technique has been fully integrated into VOSviewer. This means that VOSviewer can be used not only to view VOS maps but also to construct them. Hence, no separate computer program is needed for constructing VOS maps.

2.2.2 Vosviewer: construction of map

VOSviewer constructs a map based on a co-occurrence matrix. The construction of a map is a process that consists of three steps. In the first step, a similarity matrix is calculated based on the co-occurrence matrix. In the second step, a map is constructed by applying the VOS mapping technique to the similarity matrix. And finally, in the third step, the map is translated, rotated, and perhaps also reflected

Step 1 : Similarity matrix

VOS mapping technique requires a similarity matrix as input. A similarity matrix can be obtained from a co-occurrence matrix by normalizing the latter matrix, that is, by correcting the matrix for differences in the total number of occurrences or co-occurrences of items (Eck and Waltman, 2007b). VOSviewer, however, does not use one of these similarity measures. Instead, it uses a similarity measure known as the association strength (Eck and Waltman, 2007b; Eck *et al.*, 2006) (Van Eck and Waltman, 2007b; Van Eck et al., 2006). This similarity measure is sometimes also referred to as the proximity index (Peters and van Raan, 1993) or as the probabilistic affinity index (information *et al.*, 2004). Using the association strength, the similarity *sij* between two items *i* and *j* is calculated as

$$s_{ij} = \frac{c_{ij}}{w_i w_j}, \qquad (1)$$

Step 2: VOS mapping technique

VOS has the tendency to locate objects close to what we have called their ideal coordinates. The ideal coordinates of an object *i* are defined as a weighted average of the coordinates of all other objects, where the coordinates of objects more similar to object *i* are given higher weight in the calculation of the weighted average (Eck and Waltman, 2007c, 2009). Second, VOS seems to pay more attention to indirect similarities via third objects multidimensional scalling (MDS) (Eck and Waltman, 2007c, 2009). The idea of the VOS mapping technique is to minimize a weighted sum of the squared Euclidean distances between all pairs of items. The higher the similarity between two items, the higher the weight of their squared distance in the summation. To avoid trivial maps in which all items have the same location, the constraint is imposed that the average distance between two items must be equal to 1. In mathematical notation, the objective function to be minimized is given by

$$E(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{i<j} s_{ij} \|\mathbf{x}_i - \mathbf{x}_j\|^2, \qquad (2)$$

where the vector **x***i* = (*xi*1*; xi*2) in (2)

denotes the location of item *i* in a two-dimensional map and where $k \cdot k$ denotes the Euclidean norm. Minimization of the objective function is performed subject to the constraint in (3)

$$\frac{2}{n(n-1)} \sum_{i<j} \|\mathbf{x}_i - \mathbf{x}_j\| = 1. \tag{3}$$

Step 3: Translation, rotation, and reflection

The optimization problem discussed in step 2 does not have a unique globally optimal solution. This is because, if a solution is globally optimal, any translation, rotation, or reflection of the solution must also be globally optimal (for a discussion of this issue in the multidimensional scaling context, see Borg and Groenen, 2005). It is of course important that VOSviewer produces consistent results. The same co-occurrence matrix should therefore always yield the same map (apart from differences caused by local optima). To accomplish this, it is necessary
to transform the solution obtained for the optimization problem discussed in step 2. VOSviewer applies the following three transformations to the solution:

- **Translation** The solution is translated in such a way that it becomes centered at the origin.
- **Rotation** The solution is rotated in such a way that the variance on the horizontal dimension is maximized. This transformation is known as principal component analysis.
- **Reflection** Let *i* and *j* denote the items with, respectively, the lowest and the highest coordinate on the horizontal dimension, and let *k* and *l* denote the items with, respectively, the lowest and the highest coordinate on the vertical dimension. If $i > j$, the solution is reflected in the vertical axis. If $k > l$, the solution is reflected in the horizontal axis.

These three transformations are sufficient to ensure that VOSviewer produces consistent results.

2.2.3 Vosviewer : density map

There are two density map that include in Vosviewer which is ordinary density map and cluster density map.

Ordinary density map

In this view, items are indicated by a label in a similar way as in the label view. Each point in a map has a color that depends on the density of items at that point.

That is, the color of a point in a map depends on the number of items in the neighborhood of the point and on the importance of the neighboring items. The density view is particularly useful to get an overview of the general structure of a map and to draw attention to the most important areas in a map. The density view immediately reveals the general structure of the map. Especially the economics and management areas turn out to be important. These areas are very dense, which indicates that overall the journals in these areas receive a lot of citations. It can also be seen that there is a clear separation between the fields.

Cluster density map

This view is available only if items have been assigned to clusters. The cluster density view is similar to the ordinary density view except that the density of items is displayed separately for each cluster of items. The cluster density view is particularly useful to get an overview of the assignment of items to clusters and of theway in which clusters of items are related to each other

After long elaborate the functionality of Vosviewer, it is claimed to be the trust and robust system that can handles huge numbers of data as well as has the capabilities to creating, developing and generating the visualization of density between the words. Therefore, next task for next section is to investigate the competency and ability of Vosviewer with the real data with retrieved from the selected database.

2.2.4 Vosviewer : Output interpretation

After data been processed in Vosviewer, there are number of output that the system produced in order to categorize and list out the output.

Label.

Label is the most important output in the system Label is a list of words that being processed in the website or we can called it as ‘keywords’. These are the keywords that has the number of repetition in the publication. Some are the keywords are the normal words and has general means which every publication in every subject area used it repeatedly such as ‘articles’, ‘conference’, experiment’ or ‘papers. But some of the keywords does has some specific meaning that really shows the specific topic or area of study. Therefore, it is important to determine which words that can give a good impact to the experiment.

Cluster

According to Merriam-Webster dictionary, definition of cluster is a number of similar things that occur together such as two or more consecutive consonants or vowels in a segment of speech (Merriam-webster, 1828). Thus, in the output, column cluster has the same role which is to group the words according to their closes definition between each other. Vosviewer will help the user to identify the

number of cluster will be generated in the output, but the user also has authorize to choose how many cluster of words that they want. The bigger number of cluster, means the bigger group of words will be pooled together. From this cluster, sometimes the user can identify the potential topic or research interest that they can dig into. Vosviewer will display the cluster with difference color with difference theme as shows in figure 1.

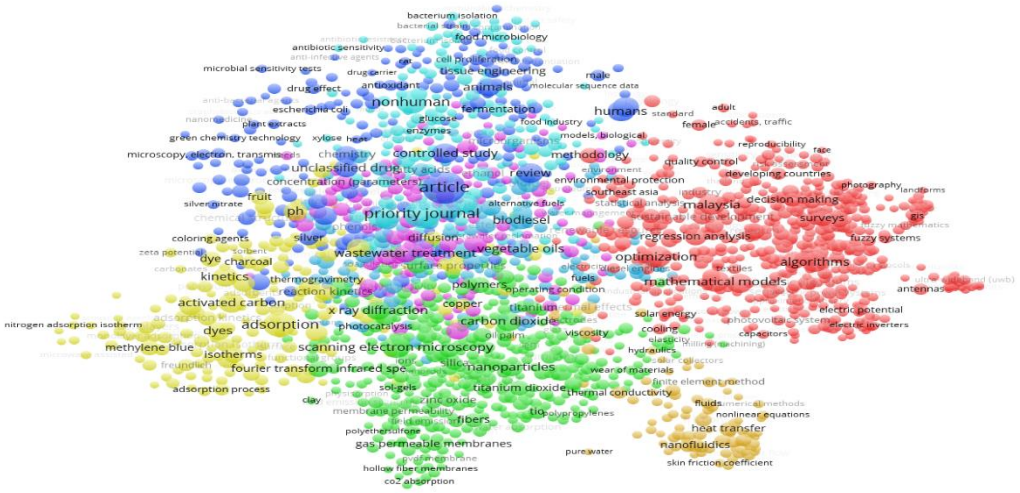

Fig 1: Cluster generated from Vosviewer

Link

Link is represent the total numbers of words that has similar meaning. In publication, words that been used by researchers to represent their words are difference with each other, but actually the words has the similar meaning or has almost the same meaning or has the same group of meaning. Thus, in figure 2, every single line which connect with the words is shows that there is another potential words that can also represent the main words.

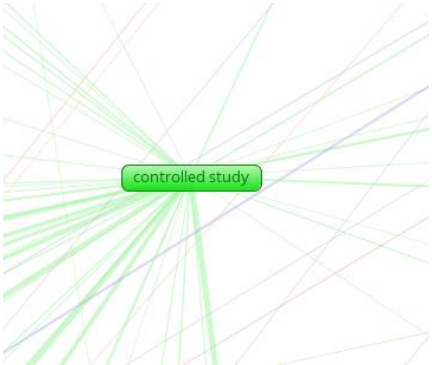

Fig 2: link of words

Occurrences

Occurrences is total number of repetition words that occurs in the data. The minimum number of occurrences that happens in the publication that been suggested by the system is 5 repetition, which means one particular words such as 'controlled study' was repeated with as minimum as 5 times in overall publication. The bigger number of repetition means that words is keep on appearing in most of the publication and shows that it is a quite number of study been made and use by researchers. Therefore, these occurrences data can gives other researchers an overview either to avoid any research topic which relates with that particular words or it might shows that it is a potential research topic that can be explores or it might has the potential to combine with other words that generated at the output that might become a new research topic. After all, every single words that generated in the output does give a significance impact in the research environment.With long elaborate in this section, next section is the experiment section with the real data which retrieved from selected database.

3 Methodology

In this application section, the implementation of stages in identify the classification of subject areas. Data that will be involved in this study within the time frame from 1995 to 2015 which is 20 years of publication by Malaysian's researchers. Data were retrieve from Scopus database.

3.1 Retrieve data from Scopus

Due to limitation in database subscription and download service for researcher's access account, not all data can be retrieved. The overall total publication in Scopus database is 294 504 for Malaysian's researchers publication but unfortunately, only 10% from the total than can be downloaded. With the recommendation by previous researchers (Glänzel, 2012), 1% of top author or top cited are selected in this study which total for to produce the co-words analysis for each subject area is between 1000 to 2000 articles. Therefore, in this study the limitation in retrieving publication data in selected database is occurred. All data were sorted from highest cited articles to the lowest cited articles. In this study total of 2360 successfully downloaded which has been sorted the highest cited articles and all subject area is selected. Thus, all articles from any subject area is sorted which gain the highest cited is included and downloaded and the distribution of publication in subject area also vary from one and another.

Reason on elect the highest cited articles is based on several factors. First factor is the age of publication and next condition is the impact from the publication towards their readers (Aksnes, 2003). Combination of both factors will helps publication gain a lot of citation. Another factors that might involves in gaining number of citation is the 'guru' in the certain subject area (Aksnes, 2003). This

happens because of the ‘guru’ it-self has their own followers or fan such as founder in certain theories or person who are discover some definition or new discoveries . Normally research and development from these ‘guru’ will give a significance impact towards nation and education, therefore, their research is important towards their readers and also can gives a direction for young researchers in developing their name in research environment.

Next task in producing the classification subject area is to clean all the ‘noise’ and uncompleted data in the database. In our downloaded data from database, almost 2% from total amount can be considered as a ‘noise’ data. The intended noise such as data isn’t in dedicated column or year of publication isn’t complete the way it should be. From the total of 2360 publication data, only 2000 can be used for the next process.

4 Analysis

All data is processed by using selected software which VosViewer. VosViewer is software that invented by Nees Jan van Eck and Ludo Waltman (Eck and Waltman, 2009) which is specialist in bibliometric mapping. The software is capable to produce the frequent keywords term that been used in title, abstract and journal title (Eck and Waltman, 2009, 2011; Eck and Waltman., 2010). One needs to determine which one should be used according to their objective of research. The reasons for VOSVIEWER is the chosen software in this study is the capabilities in handling large data (Eck and Waltman, 2007b), versatility in performing the important data (Eck and Waltman., 2010) and capability to deliver the relationship between the data which means, the closer the data, the strong relationship it have between each other (Eck and Waltman, 2009). It also capable in producing the emerging area in a subject area in selected period of time (Eck and Waltman, 2007b). VOSVIEWER also has user friendly interface which easy to interpreted compared to the others visualization bibliometric mapping software(Eck and Waltman, 2009).

Next step which is to produce cluster of publication. As being explained before, cluster of classification subject area used the co-words approach. Words from selected publication information such as title, abstract and keyword from both authors and index is processed in VosViewer. In previous research, the chosen parameter is 25. It is means that in every 25 publication which has the same words or term in each of publication, it will form a basic subject area. Unfortunately, in this study with limitation of data, we will choose parameter as a 10. Therefore, in every 10 publication, basic cluster of subject area will be formed. As for co-words or term analysis, recommendation from previous study in selecting the occurrences or repetition words in one publication is five times (Eck and Waltman, 2009). Therefore, if the words are below than five repetition will be excluded. After screening and processing experiment happens in the VOS VIEWER, there are list of potential co-words analysis is listed in the system, and

there are recommendation by the system to get only 70% from the total words. The words will be clustered according to the strength relationship between each other. Each of cluster will have one main keywords that became center of attention which mean the frequency of using the words is high in the publication.

Another important note that worth to highlight in this section, in order to generate co-occurrences and high frequency words, in Vosviewer the extraction process are involved numbers of important bibliometric element namely as title, abstract and author's keywords. Reason being is title is normally represent the overview of the articles or paper, and abstract is normally comprehensive conclusion which represent real situation which happen in the paper. Real full paper or articles can be used in generating co-words analysis but apparently, not all paper or articles can easily to be downloaded. Alternatively, only abstract, title and author's keywords can used in the software.

Thus, in this study, we will see quite big number of cluster is produced. The cluster as shows in Fig. 3.

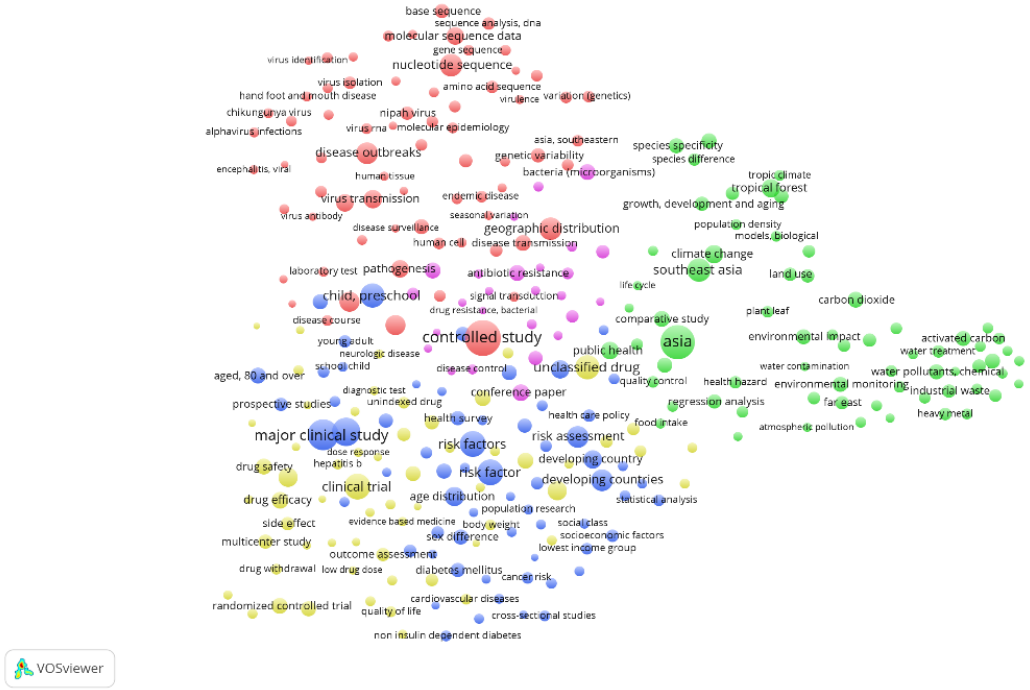

Fig. 3: Cluster generated from Vosviewer

From fig. 3, there are five cluster produced from the software. A part of details classification subject area as shows in table 1.

Table 1: A part of list co-words and cluster for classification subject area for all total publication which produce within five cluster

Cluster	Color	High frequency word	Least frequency word
Cluster 1		clinical feature, controlled study, disease outbreaks, disease severity, geographic distribution, nucleotide sequence,	acute disease, aedes albopictus, differential diagnosis, disease surveillance, encephalitis, viral gene locus, hand, foot and mouth disease human tissue, immunoglobulin m, molecular genetics, seasonal variation, sequence analysis, virulence, virus antibody, virus identification, virus replication, virus virulence,
Cluster 2		Asia, climate change, comparative study, public health, southeast asia, tropical forest,	adsorption kinetics, aqueous solutions, atmospheric pollution, central america, chemicals removal (water treatment), concentration (composition), correlation analysis, environmental exposure, food intake, life cycle, pollutant removal, polycyclic aromatic hydrocarbon, polycyclic hydrocarbons aromatic, quantitative analysis, waste disposal fluid, water contamination
Cluster 3		age distribution, child, preschool, developing countries, major clinical study, middle aged, world health	alcohol consumption, blood pressure, cancer incidence, cohort analysis, ethnic group, ethnology, health behavior, hospital admission, sex ratio, symptomatology,
Cluster 4		clinical trial, drug effect, drug efficacy, randomized controlled trial, systematic review, treatment outcome, unclassified drug	abdominal pain, antiviral agents, cardiovascular risk, confidence interval, dose response, double-blind method, drug cost, drug megadose, drug therapy, combination hepatitis c, kidney failure, low drug dose, neurologic disease, nonsteroid antiinflammatory agent, nuclear magnetic resonance imaging, patient compliance, thrombocytopenia, unspecified side effect, urinary tract infection, vitamin supplementation,
Cluster 5		antibiotic agent, antibiotic resistance, bacteria (microorganisms), conference paper, escherichia coli, infection control	amoxicillin, antibiotic sensitivity, bacterial strain, bacterial virulence, bacterium isolate, cross infection, disease control, drug resistance, bacterial, hospital infection, immune response, international cooperation, multidrug resistance, penicillin g, seroepidemiologic studies, seroprevalence, south and central america, staphylococcus aureus,

Table 2: List co-words and cluster for classification subject area for all total publication which produce within five cluster

id	label	cluster	weight<Links>	weight<Total link strength>	weight<Occurrences>
18	1,1 diphenyl 2 picrylhydrazyl	2	52	77	6
111	abdominal pain	2	258	456	11
119	abortion	5	123	184	7
124	aboveground biomass	1	54	70	7
132	absorption	1	138	167	11
138	abundance	1	128	187	12
175	acetic acid	1	111	120	6
185	acetylation	1	46	50	5
187	acetylsalicylic acid	2	311	518	14
190	aciclovir	2	155	204	7
205	acids	1	138	180	9
206	acinetobacter	5	120	210	6
214	acquired immune deficiency syndrome	5	244	347	12
225	activated carbon	1	270	634	33
251	acute disease	4	219	359	11
272	adaptation	1	235	356	17
276	adaptation, physiological	1	134	186	8
300	adhesion	1	85	101	6
314	adolescent	3	863	3617	124
334	adsorbent	1	118	201	6
338	adsorbents	1	91	140	8
339	adsorption	1	348	1226	61
341	adsorption capacities	1	98	199	11
349	adsorption isotherm	1	97	246	13
350	adsorption isotherms	1	114	346	19
351	adsorption kinetics	1	163	422	15

From the table 1 and table 2, we can identify that there are 5 cluster formed from the selected database with each of them has their own color. Each color is shows in type of cluster. Among all cluster, the most dominated words which produced from software is in medicine area which is in cluster 1, cluster 4 and cluster 5. Whereas in cluster 2, the co-occurrences words are more related to agricultural and developing a nation. In cluster 3, words occurrences are more leaning towards management, adult health and child care.

From these occurrences words generated from the selected software, we can get a knowledge that in these 20 years duration time, our researchers and scientist are more eager performed clinical, medicine and disease research. As we know, in moving towards developed country or ‘developed nation’ status, the foremost and important matter that need to put into center of focus is health of nation and overcome any disease outbreak that happens in this country. It is undeniable, among all subject area, medicine is one of area that gain a lot of publication. It is because of in medicine, there are too many division related to medicine such as drugs, clinical and others. Therefore, it is no surprise if among all cluster, medicine is dominated which is in cluster 1, cluster 4 and cluster 5.

In cluster 2, high frequency words are related to agricultural, environment and forestry. Within 20 years, there are huge development happens in order to make sure our nation get their better facilities and services. Especially in urban area, most of the facilities are been developed to avoid any deficiency such as new elevated highway, light rail transit (LRT), monorail and others. Therefore, with these development, it is indirectly affect our environment whether it is weather

changes or lack of forest. Apart from that, most of the publication revolve about Asia and Southeast Asia. This shows that, our researches are more interested to perform their research with Asia or Southeast Asia. This is also helped our government to understand better what actually happened in our country and indirectly, it can help government to plan a better infrastructure for our nation without neglect our environment.

In last cluster is cluster 3, the words generated from Vosviewer are consist about children development, adult health and health in developing countries. This cluster are somewhat the same as cluster 1, cluster 4 and cluster 5. Nonetheless, in this cluster, the focus point in more towards children health and development. These few years, development mental and physical as well as and education for children in this country become serious topic and important matter. This is because of these children will replace the current nation and they are the catalyst for future nation. Therefore, there are a lot of program, research and development that related to children development.

Nonetheless, beside high frequency words, there are also low frequency words generated from the Vosviewer. All low frequency words are shows in table 2 as well. From the result of low frequency words, there are possibilities to make advantage of it. One of the potential is to give a chance and opportunities to younger researchers to do more exploration in certain area and gives them an overview which domain research that they can involve and discover. Another potential is to acknowledge other researchers that might has potential collaboration with another research center in order to make a new discovery. From the list of words too, it can give idea for next trend research. Thus, researches need to be creative to manipulate this list of low frequency words in order to produce another publication in their area. Although in fig 1 there are certain area involves by words generated such as medicine, forestry or children management, it can give researchers from another domain to do more collaboration with these domain area.

6 Conclusions

There are few reasons on why there are only three subject area were identify in this study. First possibilities might be the most number of publication were came from these subject area. This is proven by other studies that were made from previous researchers which claimed that the most active researchers is from these subject area. As we know, in these subject area, the research and development are keep on evolve and being discovered, thus, the number of publication is keep on increasing year by year and researches from other subject area has low number of publication such as in art and humanities.

Second possibility might be the database factor. In this study, only one database which is Scopus was used to download all related publication. Scopus is infamous database which focus mainly in science and technology area. Therefore, the number of publication from other subject area especially in linguistic, literature, history and religion subject area are rather low being publish in this database. There might be difference database that mainly focus in these subject area. Thus, this is the limitation of this study which need to be improve for future study. In future study, multiple database which relate to research assessment as well as focus subject of each institutions are worth to be noted in order to get real scenario in classification subject area. The reason of not performing multiple database in this study is the limited access library subscription. In current situation, focus of institution is on technology and science, therefore, most of database subscription are mostly covered technology and science area. Therefore, only Scopus database were used in this study.

Third possible reason is classification from Scopus itself. It worth to get attention that in Scopus subject area, there are few publication falls in ‘unidentified’ subject area. We know that all subject area are keep on emerge and evolve between subject area. Nevertheless, as proven in previous study by other researchers, ‘unidentified’ in publication databases not supposedly happens in publication database as each of publication has their own genre and subject area. While retrieving related publication in Scopus, total of 483 publication aren’t fall into any subject area. This situation will be harmful and threaten the total of publication which supposedly been grouped in the specific subject area. Total publication in each subject area which be vary if the ‘unidentified’ subject area isn’t exist in the Scopus subject area.

Thus, for future recommendation, there are few point that need to give attention. First recommendation is about selection of databases. Publication from multiple databases might give another view of output cluster and will give some numbers of subject area compare to this study where there are only three subject area. Another recommendation is the total number of publication that being manipulate in order to generate the cluster. In this study, number of publication isn’t huge,

therefore, it might affect the number of cluster which generated by selected software. As in previous study, huge number of publication were manipulate in order to identify the classification of subject area. Therefore, in next study, it is necessity to gather huge number of publication is order to get the illustration of cluster as well as the classification which extract from publication.

ACKNOWLWDGEMENTS

The authors would like to thank the Ministry of Higher Education (MOHE) and the Universiti Teknologi Malaysia (UTM) for the Fundamental Research Grant Scheme (vote number: 4F621) that had supported this research.

References

[1] Ahlgren, P., Jarneving, B., and Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology Policy in the 1980s and Beyond, Longman, Harlow, 54*(6), 550–560.

[2] Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research evaluation, 12*(3).

[3] Andersen, J. P., Bogsted, M., Dybkaer, K., Mellqvist, U. H., Morgan, G. J., Goldschmidt, H., et al. (2015). Global myeloma research clusters, output, and citations: a bibliometric mapping and clustering analysis. *PLoS One, 10*(1), e0116966.

[4] Assefa, S. G., and Rorissa, A. (2013). A bibliometric mapping of the structure of STEM education using co-word analysis. *Journal of the American Society for Information Science and Technology, 64*(12), 2513-2536.

[5] Bailón-Moreno, R., Jurado-Alameda, E., and Ruiz-Baños, R. (2006). The scientific network of surfactants: structural analysis. *Journal of the American Society for Information Science and Technology*, *57*(7), 949–960

[6] Bazm, S., Kalantar, S. M., and Mirzaei, M. (2016). Bibliometric mapping and clustering analysis of Iranian papers on reproductive medicine in Scopus. *Int J Reprod Biomed (Yazd), 14*(6), 371–382.

[7] Berry, M. W. (2004). Survey of Text Mining. *Clustering, Classification, and Retrieval*.

[8] Boyack, K. W., and Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, *61*(12), 2389-2404.

[9] Braam, R. R., Moed, H. F., and Raan, A. F. J. v. (1991). Mapping of Science by Combined Co-Citation and Word Analysis I. Structural

Aspects. *Journal of the American Society for Information Science, 42*(4), 233.

[10] Buter, R. K., Noyons, E. C. M., Van Mackelenbergh, M., and Laine, T. (2006). Combining concept maps and bibliometric maps. First explorations. *Scientometrics, 66*(2).

[11] Callon, M., Courtial, J. P., and Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research the case of polymer chemistry. *Scientometrics, 22*(1), 155-205.

[12] Carley, K. M. (1997). Network text analysis: The network position of concepts. *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*, 79-100.

[13] Chen, G., Xiao, L., Hu, C. P., and Zhao, X. Q. (2015). Identifying the research focus of Library and Information Science institutions in China with institution-specific keywords. *Scientometrics, 103*(2), 707-724.

[14] Cho, C.-H., and Khang, H. (2006). The State of Internet-Related Research in Communications, Marketing, and Advertising: 1994-2003. *Journal of Advertising, 35*(3), 143-163.

[15] Coulter, N., Monarch, I., and Konda, S. (1998). Software engineering as seen through its research literature: A study in co-word analysis. *Journal of the American society for information science, 49*(13), 1206-1223.

[16] Danowski, J. A. (1993). Network analysis of message content. *Progress in communication sciences, 12*, 198-221.

[17] Dehdarirad, T., Villarroya, A., and Barrios, M. (2014). Research trends in gender differences in higher education and science: a co-word analysis. *Scientometrics, 101*(1), 273-290.

[18] Ding, Y., Chowdhury, G. G., and Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information processing & management, 37*(6), 817-842.

[19] Eck, N. J. v., and Waltman, L. (2007a). Bibliometric Mapping of the Computational Intelligence Field. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 15*(5), 625.

[20] Eck, N. J. v., and Waltman, L. (2007b). Bibliometric mapping of the computational intelligence field. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 15*(05), 625-645.

[21] Eck, N. J. v., and Waltman, L. (2007c). VOS a new method for visualizing similarities between objects. *Advances in Data Analysis. Springer Berlin Heidelberg*, 299-306.

[22] Eck, N. J. v., and Waltman, L. (2009). VOSviewer A Computer Program for Bibliometric Mapping.

[23] Eck, N. J. v., and Waltman, L. (2011). Text mining and visualization using VOSviewer. *Centre for Science and Technology Studies, Leiden University, The Netherlands, arXiv preprint arXiv:1109.2058*.

[24] Eck, N. J. v., Waltman, L., Berg, J. v. d., and Kaymak, U. (2006). Visualizing the computational intelligence field [Application Notes]. *IEEE Computational Intelligence Magazine 1*(4).

[25] Eck, N. J. v., Waltman, L., Dekker, R., and Berg, J. v. d. (2010). A Comparison of Two Techniques for Bibliometric Mapping Multidimensional Scaling and VOS. *Journal of the Association for Information Science and Technology, 61*(12), 2405-2416.

[26] Eck, N. J. v., and Waltman., L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84*(2), 523-538.

[27] Egghe, L., and Rousseau, R. (2002). Co-citation, bibliographic coupling and a characterization of lattice citation networks. *Scientometrics,, 55*(3), 349–361.

[28] Feldman, R., and Dagan, I. (1995). Knowledge Discovery in Textual Databases (KDT). *KDD'95 Proceedings of the First International Conference on Knowledge Discovery and Data Mining* 112-117

[29] Fu, H. Z., Wang, M. H., and Ho, Y. S. (2013). Mapping of drinking water research: a bibliometric analysis of research output during 1992-2011. *Sci Total Environ, 443*, 757-765.

[30] Gazni, A., Sugimoto, C. R., and Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. *Journal of the Association for Information Science and Technology, 63*(2), 323–335.

[31] Glänzel, W. (2012). Bibliometric methods for detecting and analysing emerging research topics. *El Profesional de la Informacion, 21*(2), 194-201.

[32] Glänzel, W., and Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics, 56*(3), 357–367.

[33] Glenisson, P., Glänzel, W., Janssens, F., and De Moor, B. (2005). Combining full text and bibliometric information in mapping scientific disciplines. *Information Processing & Management, 41*(6), 1548-1572.

[34] Gu, J., and Li, X. (2015). The effects of character transposition within and across words in Chinese reading. *Attention, Perception, & Psychophysics, 77*(1), 272-281.

[35] He, W., Zha, S., and Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management, 33*(3), 464-472.

[36] Healey, P., Harry Rothman, and Paul K. Hoch. (1986). An experiment in science mapping for research planning. *Research Policy, 15*(5), 233-251.

[37] Heersmink, R., van den Hoven, J., van Eck, N. J., and van den Berg, J. (2011). Bibliometric mapping of computer and information ethics. *Ethics and Information Technology, 13*(3), 241-249.

[38] Huang, M. H., Wu, L. L., and Wu, Y. C. (2015). A study of research collaboration in the pre-web and post-web stages: A coauthorship analysis

of the information systems discipline. *Journal of the Association for Information Science and Technology, 66*(4), 778-797.

[39] Igami, M. Z., Bressiani, J., and Mugnaini, R. (2014). A new model to identify the productivity of theses in terms of articles using co-word analysis. *Journal of Scientometric Research, 3*(1), 3.

[40] information, M. Z., Bassecoulard, E., and Okubo, Y. (2004). Shadows of the Past in International Cooperation: Collaboration Profiles of the Top Five Producers of Science. *Scientometrics, 47*(3).

[41] Janssens, F., Leta, J., Glänzel, W., and De Moor, B. (2006). Towards mapping library and information science. *Information Processing & Management, 42*(6), 1614-1642.

[42] Lee, B., and Jeong, Y.-I. (2008). Mapping Korea's national R&D domain of robot technology by using the co-word analysis. *Scientometrics, 77*(1).

[43] Lehmann, F. (1992). Semantic networks. *Computers & Mathematics with Applications, 23*(2-5), 1-50.

[44] Leydesdorff, L. (1996). Scientometrics and science studies from words and co-words to information and probabilistic entropy. *Journal of the International Society for Scientometrics and Informetrics JISSI 2, arXiv:1505.06007v1*

[45] Leydesdorff, L., and Welbers, K. (2011). The semantic mapping of words and co-words in contexts. *Journal of Informetrics, 5*(3), 469-475.

[46] Leydesdorff, L., and Zhou, P. (2008). Co-word analysis using the Chinese character set. *Journal of the American Society for information Science and Technology, 59*(9), 1528-1530.

[47] Li, M., and Chu, Y. (2016). Explore the research front of a specific research theme based on a novel technique of enhanced co-word analysis. *Journal of Information Science*, 016555151666191.

[48] Liu, Y., Goncalves, J., Ferreira, D., Xiao, B., Hosio, S., and Kostakos, V. (2014). Chi 1994-2013 Mapping Two Decades of Intellectual Progress through Co-word Analysis. *Proceeding CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 3553-3562.

[49] Looze, M.-A. D., and Lemarié, J. (2006). Corpus relevance through co-word analysis: An application to plant proteints. *Scientometrics, 39*(3).

[50] Merriam-webster. (1828). Definition of cluster.

[51] Muñoz-Leiva, F., Viedma-del-Jesús, M. I., Sánchez-Fernández, J., and López-Herrera, A. G. (2011). An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective. *Quality & Quantity, 46*(4), 1077-1095.

[52] Noyons, E. C. M. (1999a). Bibliometric mapping as a science policy and research management tool. *PhD Thesis. Centre for Science and Technology Studies, Leiden University, DSWO Press, Leiden.*

[53] Noyons, E. C. M. (1999b). Bibliometric mapping as a science policy and research management tool. *Doctoral Thesis*.

[54] Peters, H. P. F., and van Raan, A. F. J. (1993). Co-word-based science maps of chemical engineering. Part I: Representations by direct multidimensional scaling. *Research Policy, 22*(1), 23-45.
[55] Popping, R. (2000). Computer-assisted text analysis. *Sage*.
[56] Richard Klavans, and Boyack, K. W. (2006). Quantitative evaluation of large maps of science. *Scientometrics,, 68*(3), 475–499.
[57] Rokaya, M., Atlam, E., Fuketa, M., Dorji, T. C., and Aoe, J.-i. (2008). Ranking of field association terms using Co-word analysis. *Information Processing & Management, 44*(2), 738-755.
[58] Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science and Technology, 24*(4), 265–269
[59] Tahamtan, I., Safipour Afshar, A., and Ahamdzadeh, K. (2016). Factors affecting number of citations: a comprehensive review of the literature. *Scientometrics, 107*(3), 1195-1225.
[60] Tan, A.-H. (1990). Text Mining: The state of the art and the challenges. *Proceedings of the PAKDD 1999 Workshop on Knowledge Disocovery from Advanced Databases, 8*, 65-70).
[61] van Eck, N. J., Waltman, L., Noyons, E. C., and Buter, R. K. (2010). Automatic term identification for bibliometric mapping. *Scientometrics, 82*(3), 581-596.
[62] Viedma-Del-Jesus, M. I., Perakakis, P., Munoz, M. A., Lopez-Herrera, A. G., and Vila, J. (2011). Sketching the first 45 years of the journal Psychophysiology (1964-2008): a co-word-based analysis. *Psychophysiology, 48*(8), 1029-1036.
[63] Vlieger, E., and Leydesdorff, L. (2011). Content analysis and the measurement of meaning: The visualization of frames in collections of messages. *Public Journal of Semiotics, 3*(1), 28-50.
[64] Waltman, L., and Eck, N. J. v. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology, 63*(12), 2378–2392.
[65] Yan, E., and Ding, Y. (2012). Scholarly network similarities How bibliographic coupling networks, citation networks, co-citation networks, topical networks, coauthorship *Journal of the American Society for Information Science and Technology, 63*(7), 1313–1326